STATISTICA, anno LXIII, n. 2, 2003

THE ANALYTICAL SOLUTION TO THE PROBLEM OF STATISTICAL INDUCTION

Rodolfo de Cristofaro

1. INTRODUCTION

As is well known, induction is a kind of reasoning, which starts out from the investigation of one or more cases, and then reaches a conclusion the impact of which goes beyond the cases examined. For example, the statement "all cats like fish" is born from experimental observation of a large number of cats, none of which disproved this attitude of the feline race. Of course, one need observe only one cat that does not like fish for the result of the induction to be proven false.

In the field of statistics, induction appears in a different way, under a multiplicity of causes. For example, after noticing that the number of deaths in traffic accidents has diminished after the introduction of a driving licence with points, one can state that this reduction is due the new type of licence. But that reduction could be the result of other causes. On the other hand, in statistics a single case or a rare event (always possible, however improbable it may be) is not able to falsify the result of the induction.

The unsolved problem of induction has been held to be the scandal of philosophy ever since the day of Hume (1739). Hume's challenge to induction may be summed up in the following question: from past cases which we have had experience of, how can we draw a conclusion that goes beyond such cases? Experience, in fact, can shed light on the past, but never on the future.

The classical solutions to this problem consist of the quest for a principle legitimating the possibility of generalizations, providing an absolute assurance of their truthfulness. These include Kant's *a priori* synthetic judgement and the principle of the uniformity of nature. However, contemporary philosophy has generally recognised the impossibility of a guarantee of this kind.

In an attempt to solve the problem of induction, Popper, in his book *The Logic of Scientific Discovery* (1934), tried to reduce its dimensions. According to Popper, if the only purpose of science were that of verifying its own assertions, raising them once and for all beyond doubt, the enterprise would be desperate. The purpose of science, says Popper, is actually more limited. The scientist works by formulating hypotheses and planning experiments in an attempt to falsify his own hypotheses, rather than to verify them.

Popper calls the method he proposes *hypothetical-deductive*. However, it does not solve the problem of induction in the field of statistics, because – as we have seen – the falsification of a statistical hypothesis presents problems similar to those posed by its verification.

For interesting connections of some of the characteristics stated above, see Popper (1989), Howson and Urbach (1993).

The purpose of this article is to review the various solutions to the problem of statistical induction, with special reference to the analytical solution re-examined in the light of certain recent results.

2. INDUCTION IN THE AGE OF ENLIGHTENMENT

In the same century in which Hume wrote his treatise, two procedures of induction were proposed: Bayes' formula and Daniel Bernoulli's rule of maximizing the expected utility (called by Laplace: *moral expectation*).

As is well known, Bayes' formula allows one to update the probability of a hypothesis on the basis of available experience, while Bernoulli's rule tells us to choose – among the possible actions – the one for which the expected utility is a maximum.

Both procedures presented the same limit: the lack of a satisfactory criterion for assigning *a priori* probabilities. Actually, a principle for probability assignment did exist during the classical period (called the *principle of indifference*), which Laplace used with caution, but which fell into discredit later, when scientific positivism came to the fore, also because of the uncritical manner in which it was applied.

According to one formulation of this principle, one could assign the same probability to all the admissible hypotheses, not only when they are held to be equally plausible, but also if there is no other information and even if there is no reason to do otherwise.

From a logical point of view, it is doubtful whether ignorance can justify an assumption of equal probability for all the admissible hypotheses. This lead some to see a scandal in a principle able to produce knowledge out of ignorance, and for this reason, they rejected both Bayes' formula and the rule of expected utility. In a way, this meant throwing the baby out with the bathwater, thus leading to a long period of crisis.

In particular, Boole (1854) saw the *a priori* probabilities as arbitrary constants, and this led him to consider a precise solution of the problem of statistical inference to be impossible. Gini (1939) confirmed this point of view: if we do not have the necessary knowledge concerning prior probabilities – and practically, except in particular cases, we do not have – it is absurd to pretend to calculate posterior probabilities.

3. FISHER'S INDUCTIVE REASONING

It was in this crisis that Fisher focused his attention on the *significance tests*. Bernoulli and Laplace had already used these tests as analytical indicators. However, Fisher used them as instruments of inference (or as methods of induction by elimination).

The way in which Fisher set up his theory is associated with his experience in the design of experiments. In this context, the interpretation of observed data is the result of a challenge launched by the experimenter against the changing play of chance in the observation of experimental reality. The hypothesis which he set up to test (called the *null hypothesis*) is that the result of the experiment is due to chance (accidental causes).

According to Fisher, the important element is that the experimenter must be willing to run the risk of rejecting a hypothesis that could turn out to be really true. Besides, this risk must be established before performing the experiment.

With Fisher, induction is reduced to a test of the null hypothesis, with a dubious dichotomy: either the result of the experiment depends solely on chance, either it is due to an experimental cause. Quoting Fisher's own words (1956, p. 39): *Either* an exceptionally rare chance has occurred, *or* the theory of random distribution is not true".

Fisher's inductive reasoning poses several problems, from both a logical and a practical point of view. In this connection, see Howson and Urbach (1997). Besides, it is worthwhile noticing that the hypothesis of strict randomness is difficult to realise in practice. This means that with, a sufficiently large sample, a small difference from the null hypothesis may be significant, while, on the contrary, a large difference may not be significant. As has been noticed, an insignificant result, far from telling us that the effect is non-existent, merely warns us that the sample was not large enough to reveal it.

Despite this, many attempts have been made to justify significance tests. One attempt uses the Bayesian intervals of confidence, and involves rejecting the hypothesis $\theta = \theta_0$ at level α if θ_0 is not included in the smallest estimate interval θ with probability 1-α. See, for example, Lindley (1965), Box and Tiao (1973). This attempt however is still dubious, since the argument used contains a logical leap.

The fact that the null hypothesis is or is not included in the confidence interval does not mean it is true or false, nor does it mean that it should be accepted or rejected. All the hypotheses within the interval are credible, and there is no reason to privilege a single hypothesis.

Another attempt consists of replacing the significance judgement with the *p-value* (also known as the *observed significance level*). Actually, knowing the *p-value* does not solve the typical problems of the test, especially if the experiment has not been planned in a satisfactory manner.

Concerning this matter, one can remember the application of the chi-square test made by Fisher (1970) to Weldon's data (concerning the frequency distribution of the number of dice with a score higher than four out of 26,306 throws of 12 dice). In the case under examination, the *p-value* $\cong$ 0.0001. The difference between theory and experience is significant beyond any usual significance level, although the defect existing in that cast of dice is practically negligible and inevitable in the performance of any random experiment.

Knowledge of the *p-value* does not change the heart of the problem in other applications, such as – for example – variance analysis. For a proper analysis of variance, it should be remembered that the posterior distribution of variance between samples was found by Lecoutre (1984, 1985), who called it psi-square.

For another recent attempt to re-propose a new significance test, see Pereira and Stern (1999).

4. INDUCTION ACCORDING TO NEYMAN-PEARSON THEORY

Neyman has the merit of having noted that there is no logical implication connecting the observation of experimental data to the rejection of a hypothesis. This means that the rejection of the null hypothesis is not the consequence of logical reasoning, and hence of an induction in the proper sense of the word, but rather of an act of will power.

This is why, ever since the thirties, Neyman introduced the notion of inductive behaviour, thus contributing to transforming the problem of induction into a problem of decision-making. According to Neyman (1957), accepting a hypothesis does not mean stating that it is true, but only that it is convenient to behave as if it were true. What a matter is to cut down the risk of error: on the one hand, when rejecting a true hypothesis, on the other hand, when accepting a false hypothesis.

Neyman and Pearson (1933) formulated a theory for testing hypotheses where the inductive problem was dealt with as an optimum problem, in the sense of leading to success in the highest possible number of applications.

Later, Wald (1950) developed Neyman's ideas within an explicit decision theory, using the results of another theory proposed some years before by J. von Neumann, the game theory. Now, a principle used in choosing strategies developed in game theory is that of *minimax,* which Wald referred to in order to make up for lack of knowledge of *a priori* probabilities.

In any case, the community of statisticians was struck by a result achieved by Wald, that is that any decision rule which violates Bayes' formula, is inadmissible, in the sense that is dominated by another rule, consistent with the axioms of probability, uniformly better than the former. Although these results do not indicate how prior probabilities should be assigned, they do show that a reasonable behaviour is equivalent to their assignment and vice versa.

Despite this revaluation of the canonical procedure of inference, in statistics courses, large room is still devoted to the tests of hypotheses. Actually, such tests may find justification within the context of the decision theory, at least for a certain class of hypotheses (known as simple hypotheses), as a special case of the minimax rule. However, it should be remembered that minimax rule was put into discussion by Lindley (1985).

On the other hand, it is worth noticing that if the test of hypotheses is applied to data, which have already been observed, it basically still has the same problems posed by the relevant significance test.

To my way of thinking, Neyman-Pearson's theory has survived through time because of the need – strongly felt by the scientific community – for an objective solution to the problem of induction (or, as we shall see below, of an analytical solution).

5. THE SUBJECTIVE SOLUTION

In the subjective approach, inductive reasoning is the mental process according to which our 'opinions' must be modified on the basis of the data. According to de Finetti (1937), "it is thus that when the subjective point of view is adopted, the problem of induction receives an answer which is naturally subjective but in itself perfectly logical."

Actually, from a pragmatic point of view, de Finetti's theory may be seen as a possible solution to the problem of induction. However, we can still ask ourselves: is it the answer to the problem of induction as posed by Hume? It is hard to answer this question in the affirmative.

Passing on the justification of induction to the subjective responsibility of the researcher is trivial, and maybe this is not what Hume was asking for. On the other hand, the inductive process, though not mechanical or automatic, should be controlled analytically all along its track; otherwise it is only partially logical.

In any case, there are certain novelties in de Finetti's approach that make it come quite close to a logical solution in the analytical sense.

One novelty of de Finetti's approach involves the introduction of the *principle of coherence*. As is well known, a probability evaluation is called coherent if it does not leave the possibility of working out a combination of bets able to ensure a sure gain (or a sure loss).

If we accept this principle, then the usual probability axioms and the consequent Bayesian procedure appear inevitable in inductive reasoning. A reasoning that is founded on a probability interpreted as the price a person is willing to pay in a bet. Of course, the evaluation of this probability appears like the evaluation of any other price, which is made in a process of subjective synthesis.

Another novelty concerns the hypothesis of independence. In particular, this hypothesis can be assumed conditionally to the knowledge of a parameter. However, if such a parameter is not known, we cannot formulate a hypothesis of independence, otherwise we would have nothing to learn from experience. This is why de Finetti suggested replacing the condition of independence with the less restrictive condition of *exchangeability*, where the assignment of probabilities is indifferent to the order in which the events take place.

De Finetti drew his inspiration from the pragmatism of Giovanni Vailati (1863-1909) and, indeed, we can see his as a pragmatic kind of solution, although it does not necessarily turn into a decision. According to de Finetti (1977), the moment of statistical induction results as autonomous due to the peculiarity of its function, as a linking element between the knowledge we start from and consequent choice of the decision proved as the most suitable. This linking element, which responds to one of its specific and well-defined functions, consists in the application of Bayes' theorem.

6. INDUCTION SEEN AS A DECISION PROBLEM

Another pragmatic kind of solution, which has helped to revaluate the subjective solution, is that of Savage, where induction is seen as a problem of decision making.

Savage's decision theory takes into account a previous result obtained by Ramsey. In particular, in order to justify the probabilities used in betting, Ramsey (1931) introduced axioms of consistence on the options and showed that, under such axioms, there exists a function of utility and a value of probability (in the sense of *degree of belief*).

At this point, all the conditions were there for proposing the Bernoulli's rule of decision again. Savage's merit (1954) was that of having provided a justification for it within a well-defined system of axioms. For this reason, he is rightly considered the Euclid of Statistics (cf. Lindley, 1990).

In other words, Savage showed that the model of maximisation of moral expectation is merely a consequence of some simple and natural conditions on the order of preferences of the person acting. Savage also demonstrated how, under such conditions, there exists, and is also unique, both a function of probability and a function of utility for the person who takes a decision (in this sense, subjective).

7. THE ANALYTICAL SOLUTION

According to Carnap (1962), the problem of determining the degree of confirmation of a hypothesis h on the basis of the evidence e is not of an empirical nature, but is a problem of a logical kind, which can be answered analytically.

On the contrary, decision theory is the outcome of a synthesis where the subjective component escapes any logical analysis. By this, we do not intend to deny the presence of various sources of information, but the induction analyst must know how far the induction is determined by the personal knowledge of the researcher, and what the result of the induction would be in the case of irrelevance of other information.

Carnap's analytical solution is tied to a revaluation of the original notion of inference. In particular, inference is not a synonym for induction. Inference is a logical procedure through which one reaches a conclusion starting out from statements assumed as premises.

In the theory developed by Carnap inductive inference – though leading to various possible alternatives, each of which is associated with a probability value

– has a unique and rigorous character that, from this point of view, does not differ from deductive inference.

Carnap himself distinguishes among three types of inductive inference (which he places on the same logical level): 1. *direct* (the inference from the population to a sample); 2. *inverse* (the inference from a sample to the population); 3. *predictive* (the inference from one sample to another sample not overlapping with the first).

According to Carnap, many of the methods of mathematical statistics are usually not based on a system of inductive logic, but developed independently. Similarly, deductive mathematics was first developed independently of logic for more than two thousand years. Inductive logic, if sufficiently developed, will serve as a logical foundation for the methods of mathematical statistics with a clarification of the foundations of induction, of the presuppositions of induction, which are hardly ever made explicit, and the meaning and conditions of its validity.

Thus the task and function of inductive logic is analogous and complementary to that of deductive logic (Carnap, 1962, p. 222). This notion is shared by Jaynes (2003): as a matter of fact, no antithesis exists between deduction and induction. There is a single inference machine, whose kernel is *ordinary logic*. Likewise, according to Howson and Urbach (1993), Bayesian rule is analogous in certain respects to the deductive rule of *modus ponens*.

Of course, inductive inference must reach conclusions that are consistent with the axioms, or with choices that are suitable for probability theory. If the inference cannot be deduced from the axioms, we will no longer be able to provide a rational justification for it. Carnap (1971) in fact noted that accepting or rejecting hypotheses is a pragmatic choice, which does not belong to inductive logic; into which it would again risk falling under Hume's criticism.

As soon as the axioms of a theory are accepted, the latter becomes self-sufficient. The truth of its results is relative to our conventions. We can discuss these conventions, but we are not allowed to abandon the axioms halfway along and set up the inference on the basis of new rules.

The situation is no different in statistical inference. If we do not like the axioms of probability, we can always change them, but we cannot invent methods of inference irrespective of the axioms, or make choices beyond the purposes of probability theory.

In other words, inductive inference is uniquely determined by the axioms of probability. Under such axioms, statistical inference correctly describes the inductive situation and (to the contrary of induction) cannot be disproved or invalidated by new evidence.

As Carnap commented (1962, p. 518), the methods developed by Fisher, Neyman, Pearson, and Wald or new methods of a similar nature are ingenious devices, but they do not belong to inductive logic. Of course, the same argument is valid for the recent full Bayesian significance test. See, with regard to this, Madruga, Pereira and Stern (2003).

To be more explicit, any statistical procedure that is not derived from the axioms of probability theory and does not use the language of that theory is not technically an inference. The conclusion is that statistical inference is firmly based on probability alone (cf. Lindley, 2000).

As you can read in the preface to Carnap's book (1962, p. xv): "In contrast to the customary view that the outcome of a process of inductive reasoning about a hypothesis *h* on the basis of given evidence *e* consists in the acceptance (or the rejection, or the temporary suspension) of *h*, I believe that the outcome should rather be the finding of the numerical value of the probability of *h* on *e*."

To conclude, inferring, in scientific induction, means assigning a probability value (*degree of confirmation*) to hypotheses on the bases of the evidence within a well-defined system of axioms.

8. PROBABILITY IN INDUCTIVE INFERENCE

The analytical solution to induction requires a notion of probability taken as the logical relationship between a hypothesis and the evidence; real evidence, not hypothetical evidence. From this point of view, the probability of an event cannot be interpreted, in the manner of von Mises, as frequency-limit. This interpretation is not factual, and rules out any possibility of learning from experience.

As we observe new cases, experience changes and probability evaluation changes with it. In other words, the probability may not be referred to an experience that has not yet taken place. In this sense, the interpretation being discussed is not correct. In any case, it is not suitable for inductive inference.

Referring to the Bernoulli model of lottery, the probability may be interpreted as frequency in the urn or as a forecast of a frequency in the long run. The empirical law of randomness and the law of large numbers justify these interpretations of probability. However, any probability, ever since it is used in inductive inference, is a logical probability, irrespective of the real course of events.

According to Carnap (1962, p. 188), many writers since the classic period have said of certain probability statements that they are 'based on frequencies' or 'derived from frequencies'. Nevertheless, these statements often, and practically always if made before the time of Venn, speak of logical probability. They are (logical) probability statements referring to an evidence involving frequencies.

To tell the truth, difficulties still remain in evaluating probability before any factual knowledge is available (the so-called *null confirmation*) or when prior knowledge is held to be irrelevant. From our point of view, however, these difficulties arise because we have not properly clarified the elements such an evaluation depends on.

According to Carnap, the numerical value of this purely logical function of *h* only depends on the meaning of *h*. In other words, on its field as this is determined by the semantic rules of language to which *h* belongs.

Unfortunately, this formal part of Carnap's work is not very convincing, and has not indeed met with much consent in the community of statisticians. Jeffreys' principle of invariance (1961) has met with much more success, as has the criterion he obtained from this principle (based on Fisher's measure of information).

Box and Tiao (1973) introduced a similar criterion. According to these authors, the prior distribution for θ is assumed to be locally uniform, if different sets of data translate the likelihood curve on the θ-axis, leaving it otherwise unchanged (that is, the data only serve to change the location of the likelihood). On the other hand, if θ is not *data translated*, Box and Tiao suggest expressing the parameter θ in terms of a new metric $\phi(\theta)$, so that the corresponding likelihood is data translated.

The translation criterion met with a short season of interest in the early seventies. Later, it fell into discredit (at least from a logical point of view), since it takes into account the data collection method, thus violating the likelihood principle.

However, as we previously showed (1996, 2004), the likelihood principle is not a direct consequence of Bayes' formula. This is a counter-example about likelihood principle. In a referendum for the final rejection of a law, the counting of votes goes on in all constituencies so far as r (1 or more) consecutive pros are observed. Suppose that in such a way the vote supports the law. In view of likelihood principle this stopping rule should not affect the inference. But this view appears doubtful and it should be hardly accepted from the referendum promoters. In this connection, for $r = 1$, see (5. 1) in de Cristofaro (2004) and the relative example for $\theta = 1/3$.

In reality, information about design is one part of the *evidence* and is relevant for the prior. Let us see this point in a simple way. As is well know, when a coin is tossed, heads and fail have the same probability if the coin is fair. Likewise, apart from other information, we will assign the same probability to all the possible hypotheses if the sampling is *fair* or *impartial*.

To be precise, the equiprobability assumption is linked to the idea of the impartiality of design with respect to the parameter under consideration. As can be seen from the examples shown by Box and Tiao (quoted, pages 27-39), and as we have previously supported, this impartiality happens if the likelihood curve relative to that parameter, besides being symmetrical, is also *data translated*.

As Carnap demanded (1962, p. 518), why did statisticians spend so much effort in developing methods of inference independent of the probability axioms? It seems clear that the main reason was purely negative; it was the dissatisfaction with the principle of indifference (or insufficient reason). If we should find a degree of confirmation which does not lead to the unacceptable consequences of the principle of indifference, then the main reason for developing independent methods of estimation and testing would vanish. Then it would seem more natural to take the degree of confirmation as the basic concept for all of inductive statistics.

Today, the unacceptable consequences of the principle of indifference may be avoided through the notion of impartiality of the design used in inference. Therefore, statisticians would now support Bayes' rule as the best instrument for inverse inference (Notice that Fisher, Neyman, and Wald did not deny Bayesian approach in case a prior was available).

This much having been said, the analytical problem of induction can be said to have been solved: the conclusion of a research can be notified in intelligible form (not only when prior knowledge is held to be irrelevant), making all the sources of information that determined it explicit.

9. PRAGMATIC OR ANALYTIC SOLUTION?

The violation of the axioms of the utility theory has been compared to a stupid mistake, such as that committed by a person who tries to trace more than one straight line between two points. However, even the most obvious axioms may have implications that are not easy to imagine on first sight. For example, if the probabilities exist, this means that their existence is already implied by the axioms.

According to some scholars, demonstrating the existence of the function of utility would force researchers to give it a numerical evaluation. According to others, one thing is to demonstrate the existence of the function of utility, another is to measure it in practice, and as long as one does not know how to measure it, decision theory is a difficult theory to apply.

For instance, in game theory, it is proved that, for every kind of game with complete information – such as chess – there is an optimum strategy, which should not be given up by either party. However, the number of possible moves in a game of chess is so high that this strategy has not yet been found (luckily for all those who love chess).

Lindley (1985) acknowledges that most people are not logical and do not reason in terms of utility and probability. However, he believes it is possible to teach them how to reason correctly.

In decision theory, the possibility that one utility (or a loss) can be infinite is ruled out. In practice, however, we can find ourselves dealing with situations having unlimited utilities. For example, if we admit that human life has no price, the discovery of a drug able to cure a mortal illness would have an unlimited utility. On the other hand, the utility of scientific hypotheses is an epistemic utility, difficult to assess in monetary or quantitative terms.

Other discussions concern the actual behaviour of individuals as opposed to the behaviour to be expected on the basis of Savage's independence axiom (or principle of the sure thing). Observations of decision making of single human beings have shown the existence of frequent violations of Savage's independence axiom (the Allais and Ellesberg paradoxes). It is suggested that these violations result from the fact that under some situations of choice under uncertainty or of partial ignorance the independence axiom is necessarily violated by rational decision makers. Thus the claim to universal observance of this axiom is denied. Regarding this issue, see Cohen and Jaffray (1988) and the references listed there.

On the contrary, according to the supporters of decision theory, the criterion of the expected utility is an obligatory path because, if we follow a different criterion, we are forced to behave in an absurd manner, for example going towards a sure loss.

Some uphold the view that decision theory is more complete compared to inductive logic, which does not consider the cost of experimentation, the operative value of the conclusions reached and their consequences. Some go further, upholding the view that any conclusion which a researcher may draw from the data observed must be operative, otherwise the same conclusion is not such, or must be considered to be meaningless. Regarding this, see Smith (1984).

The need to keep the two solutions separate can easily be understood if we examine certain situations where the application of decision theory does not appear to be very satisfactory from an analytical point of view. This topic will be briefly mentioned in the following points:

1. Acceptance of a hypothesis is an operative conclusion that must be kept separate from its credibility. For example, the decision to insure one's house against fire does not mean that the hypothesis of fire is very probable. The error may appear to be banal, but it is not rare, in statistics, to find what is only more convenient considered as more probable.

2. Decision theory only takes into account the expected utility. If the expected utility of an action is even slightly greater than that of another, the first is the action that the individual must prefer, regardless of the risk involved in his decision. For example, in Italy, because of the constant tax amnesties, it may be convenient not to pay taxes. However, the businessman who does not pay taxes must be aware that he risks bankruptcy and, in some cases, going to gaol.

In a certain sense, decision theory is a regression, for statistics, to its lowest level, back to when we merely stopped at examining the average: especially when, at the beginning of the 20th Century, people used to say – with a touch of humour – that, according to statistics, all men die at the age of 40 (at the time, this was the lifespan).

3. The axioms of decision theory imply that a decision must be made, whatever the information may be. Now if we do not add to the possible actions that of postponing the decision to the time when further information will be available, some problems may arise. For example, you can decide to give up a hypothesis about the efficacy of a drug on the basis of a small sample or a partial experiment.

4. Scientific research goes on through time. From this point of view, it appears that it is more suitable to have a theory, which provides the probability of a hypothesis, and updates it from time to time as new experiments are carried out and new data are gathered, compared to the replies, which other theories can offer.

5. Decision theory offers an optimum solution with reference to a specific working context. As Barnett (1982, p. 226) noted: "Any decision rule that is inadmissible is considered to be so *with respect to the particular loss function, which is being used*." In other contexts (or if the context has not been analysed properly), the conclusion of a decisional procedure may be somewhat slanted.

As a matter of fact, inductive logic allows us to provide a thorough description of the inductive situation, thus correcting the distortions, which may arise due to an uncritical application of other theories. In other words, inductive inference is a thing; other thing is its consequence in an operative context, with a different system of axioms, and a particular loss function.

To tell the truth, both solutions (pragmatic and analytic) are legitimate and autonomous within their respective axioms. Of course, any other approach deserves to be examined in view of possible further solutions to the inductive problem, as long as the axioms it is based on are made explicit.

Dipartimento di Statistica "Giuseppe Parenti" RODOLFO DE CRISTOFARO
Università di Firenze

REFERENCES

V. BARNETT (1982), *Comparative Statistical Inference.* 2nd ed. John Wiley and Sons Ltd.

D. BERNOULLI (1730-31) *Specimen theoriae novae de mensura sortis*, "Commentarii Acad. Sc. Imperialis Petropolitanae", vol. 5, 175-192. English translation on "Econometrica", vol. 22, 23-36.

G. BOOLE (1854), *An Investigation of the Laws of Thought.* Walton and Maberley, London.

G.E.P. BOX, G.C. TIAO (1973), *Bayesian Inference in Statistical Analysis.* Addison Wesley, California.

R. CARNAP (1962), *Logical Foundation of Probability*, 2nd ed. University Press, Chicago.

R. CARNAP (1971), *Inductive logic and rational decisions.* "Studies in Inductive Logic and Probability", vol. 1, Carnap R., Jeffrey R.C. Eds. Berkeley.

M. COHEN, J. Y. JAFFRAY (1988), *Is Savage's independence axiom a universal rationality principle?* "Behavioral Science", vol. 33, pp. 38-47.

R. DE CRISTOFARO (1996), *L'influence du plan d'échantillonnage dans l'inférence statistique*, "Journal de la Société de Statistique de Paris", 137 (4), 23-34.

R. DE CRISTOFARO (2004), *On the foundations of likelihood principle*, "Journal of Statistical Planning and Inference" (Article in print).

B. DE FINETTI (1937), *La prévision: ses lois logiques, ses sources subjectives*, "Annales de l'institut Henri Poincaré", 7, 1-68.

B. DE FINETTI (1977), *Bayesian statistical inference*, translation of the paper published in the Proceedings of the Congress on "I fondamenti dell'inferenza statistica" held in Florence on the April 1977, in "Statistica", Anno LII, Supplemento al n. 3, 513-524, 1992.

R. A. FISHER (1970), *Statistical Methods for Research Workers*, 14th edition. Oliver & Boyd, Edinburgh. (First published 1925).

R. A. FISHER (1956), *Statistical Methods and Statistical Inference,* Oliver & Boyd, Edinburgh.

C. GINI (1939), *I pericoli della statistica*, Opening lecture at the I Riunione Scientifica della Società Italiana di Statistica, Pisa, 9 Ottobre 1939. English translation: *The dangers of Statistics*, "Statistica", Anno LXI, Supplemento al n. 1, 257-299, 2001, Bologna.

C. HOWSON, P. URBACH (1993), *Scientific Reasoning: The Bayesian Approach*, 2nd. ed., Open Court, Chicago and La Salle, Illinois.

D. HUME (1739), A treatise of Humane Nature, London.

E. T. JAYNES (2003), *Probability Theory: The Logic of Science.* Edited by G. Larry Bretthorst. Cambridge University Press.

H. JEFFREYS (1961), *Theory of Probability*, 3rd ed. Clarendon Press, Oxford.

B. LECOUTRE (1984), *L'analyse Bayesienne des comparaisons*, Presses Universitaires de Lille.

B. LECOUTRE (1985), Reconsideration of the F-test of the analysis of variance: the semi-bayesian significance test. "Communications in Statistics: Theory and Methods", 14 (10), 2437-2446.

D. V. LINDLEY (1965), Introduction to Probability and Statistics from a Bayesian Viewpoint, vol. 2. University Press, Cambridge.

D. V. LINDLEY (1985), *Making Decisions*, 2nd ed. John Wiley and Sons Ltd.

D. V. LINDLEY (1990) The 1988 Wald Memorial Lectures: The present position in Bayesian statistics. "Statistical Science", 5 (1), 44-89.
D. V. LINDLEY (2000), *The philosophy of statistics*, "The Statistician", 49, Part 3, 293-337.
M. R. MADRUGA, C. A. B. PEREIRA, J. M. STERN (2003), *Bayesian evidence test for precise hypotheses*, "Journal of Statistical Planning and Inference", 117, n. 2 185-198.
J. NEYMAN (1957), *'Inductive behavior' as a basic concept of philosophy of science*, "Revue de l'Institut International de Statistique", 25, 7-22.
J. NEYMAN, E. PEARSON (1933), *On the problem of the most efficient tests of statistical hypotheses*, "Philosophical Transaction of the Royal Society", Series A, 231, 289-337.
C. A. B. PEREIRA, J. M. STERN (1999), Evidence and credibility: full Bayesian significance test for precise hypotheses. "Entropy", 1, 69-80.
K. R. POPPER (1989), *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, London.
F. P. RAMSEY (1931), The Foundations of Mathematics and other Logical Essays, Adams and Co., New Jersey.
L. J. SAVAGE (1954), *The Foundations of Statistics*, John Wiley and Sons Ltd.
A. F. M. SMITH (1984), *Bayesian Statistics. Present Position and Potential Developments: Some Personal Views*, "Journal of the Royal Statistical Society", Series A, 147, Part 2, pp. 245-259.
A. WALD (1950), *Statistical Decision Functions*, John Wiley and Sons Ltd.

RIASSUNTO

La soluzione analitica al problema dell'induzione statistica

Questo articolo è una personale rassegna della storia e dei fondamenti dell'induzione statistica. Le principali impostazioni proposte per risolvere questo problema sono state esaminate in relazione ai loro meriti, con particolare riferimento all'impostazione analitica sostenuta da Carnap. Questa soluzione è stata riesaminata in base ad alcuni recenti risultati ed è stata nuovamente proposta all'attenzione della comunità degli statistici.

SUMMARY

The analytical solution to the problem of statistical induction

This article is a somewhat personal review of the history and substance of the problem of statistical induction. The main approaches proposed for solving this problem have been examined according to their merits, with special reference to the analytical solution supported by Carnap. This solution has been re-examined in view of certain results, and is proposed again to the attention of the statisticians.